

УДК 004.89

В.В. Яцишин канд. техн. наук, доц., Р.І. Капаціла

Тернопільський національний технічний університет імені Івана Пулюя, Україна

АНАЛІЗ СУЧАСНИХ АЛГОРИТМІВ ОПТИЧНОГО РОЗПІЗНАВАННЯ ОБРАЗІВ

V.V. Yatsyshyn Ph.D., Assoc. Prof., R.I. Kapatsila

ANALYSIS OF MODERN OPTICAL PATTERN RECOGNITION ALGORITHMS

Питання розпізнавання образів протягом багатьох десятиліть залишається актуальною задачею, що охоплює велику кількість теоретичних та практичних реалізацій. На сьогодні існує ряд алгоритмів, програмні реалізації яких успішно використовуються у реальних проектах.

Залежно від наявності або відсутності прецедентної інформації, розрізняють задачі розпізнавання з навчанням і без навчання. Задачі розпізнавання на основі наявної множини прецедентів називають класифікацією з навчанням (або з учителем) [Літ 1]. У випадку, коли існує множина векторів ознак, отриманих для деякого набору образів, але правильна класифікація цих образів невідома, виникає задача поділу цих образів на класи за подібністю відповідних векторів ознак. Цю задачу називають кластеризацією або розпізнаванням без навчання. Проведемо аналіз найбільш використовуваних на практиці підходів щодо оптичного розпізнавання образів.

Одним з класичних підходів до розпізнавання образів є підхід, який базується на використанні байєсівських мереж. Байєсівський підхід передбачає моделювання образів з точки зору статистичної природи спостережень [Літ 2]. За основу береться припущення про існування ймовірнісної міри у просторі образів, яка або відома, або може бути оцінена. За мету, в даному випадку, ставиться розробка такого класифікатора, який буде правильно визначати найбільш ймовірний клас для тестового образу. Тоді задачі розпізнавання образів із застосуванням Байєсівського підходу полягають у визначенні "найбільш ймовірного" класу. Постановка задачі у формальному вигляді має наступний вигляд. Нехай задано M класів $\Omega_1, \Omega_2, \dots, \Omega_M$, а також $P(\Omega_i|x)$, $i = 1, 2$ – ймовірність того, що невідомий образ, який представляють вектором ознак x , належить класу Ω_i .

$P(\Omega_i|x)$ називається апостеріорною ймовірністю, оскільки задає розподіл індексу класу після експерименту (а posteriori - тобто після того, як значення вектора ознак x було отримано). Недоліком цього методу є необхідність чіткого вказання існування апіорного розподілу для невідомого параметра, так і знання його форми.

Для розпізнавання зображень також можна застосовувати алгоритм перцептрона. В алгоритмі перцептрона в основу покладено принцип дії нейрона. Даний алгоритм представляє собою послідовну ітераційну процедуру [Літ 3]. Кожен крок полягає в пред'явленні нейрону чергового вектора-прецеденту і корекції ваг W_i за результатами класифікації. При цьому прецеденти визначаються циклічно, тобто після визначення останнього знову пред'являється перший. Процес навчання закінчується, коли нейрон правильно класифікує всі прецеденти.

При використанні схеми Кеслера, ідея побудови лінійного класифікатора природно узагальнюється на випадок класифікації з числом класів більше двох. Розглянемо задачу класифікації по M класах. Для кожного класу необхідно визначити лінійну дискримінантну функцію W_i , $i = 1, 2, \dots, M$. Нехай $x = (l + 1)$ – мірний вектор в розширеному просторі.

Вектор x відноситься до класу Ω_i , якщо $W_i x > W_j x$, $\forall i \neq j$.

Схема Кеслера дозволяє застосувати алгоритм персептрона для вирішення цього завдання, і в результаті задача розпізнавання оптичних образів полягає у побудові лінійного класифікатора в $(l+1)M$ – вимірному просторі так, щоб кожен з $(M-1)N$ – векторів-прецедентів лежав позитивному напівпростору. Якщо вектори у вихідній задачі можна розділити, то при цьому ефективним є застосування алгоритму персептрона.

Також для класифікації зображень та підвищення якості результатів можна використовувати і багатошаровий персептрон.

Кожен елемент мережі будує зважену суму своїх входів з поправкою у вигляді доданка і потім пропускає цю величину активації через передавальну функцію, і таким чином обраховується вихідне значення цього елемента. Елементи організовані в пошарову топологію з прямою передачею сигналу. Таку мережу легко можна інтерпретувати як модель вхід-вихід, в якій ваги і порогові значення (зміщення) є вільними параметрами моделі. Така мережа може моделювати функцію практично будь-якого ступеня складності, причому кількість шарів і кількість елементів в кожному шарі визначають складність функції. Визначення кількості проміжних шарів і кількості елементів в них є важливим питанням при конструюванні багатошарового персептрону. Оптимальну кількість нейронів у шарі та самих шарів знаходять експериментальним шляхом.

Кількість вхідних і вихідних елементів визначається умовами задачі. В якості початкового наближення можна взяти один проміжний шар, а кількість елементів у ньому обрати рівним півсумі числа вхідних і вихідних елементів [Літ 4].

Після того, як визначено кількість шарів і кількість елементів в кожному з них, потрібно знайти значення для ваг і порогів мережі, які б мінімізували помилку прогнозу, що видається мережею. Саме для цього служать алгоритми навчання. З використанням зібраних статистичних даних ваг і порогові значення автоматично коригуються з метою мінімізувати це відхилення. По суті цей процес представляє собою «підгонку» моделі, яка реалізується мережею, до наявних навчальних даних. Помилка для конкретної конфігурації мережі визначається шляхом прогону через мережу всіх наявних спостережень і порівняння реальних вихідних значень з бажаними (цільовими) значеннями. Всі такі різниці підсумовуються в так звану функцію помилок, значення якої і виражає помилку мережі. В якості опції помилок найчастіше береться сума квадратів помилок, тобто коли всі помилки вихідних елементів для всіх спостережень зводяться в квадрат і потім сумуються. В цілому даний метод є доволі ефективним, тому часто застосовується у сучасних розробках.

У даній роботі розглянуто лише найбільш загальні принципи побудови систем розпізнавання та класифікації зображень. В цілому існує куди більше алгоритмів розпізнавання, і кожна із систем може використовувати власний підхід у залежності від поставленого завдання.

Література

1. Потапов А.С. Распознавание образов и машинное восприятие. - С-Пб.: Политехника, 2007. - 548 с.
2. Минский М., Пейперт С. Персептроны. - М.: Мир, 2007. - 261 с.
3. Джордж Стокман, Линда Шапиро [Computer Vision Компьютерное зрение]. — М.: Бином. Лаборатория знаний, 2006. — 752 с.
4. Уздин Д. Новые меры близости, функции состояний и решающие правила в теории распознавания образов (состояний). – М.: МАКС Пресс, 2015. – 96 с.